



TOP PAGE | QUERY | RESULTS | PROJECTS | VIEWS | DATABANKS

**HELP****Reset****View**

\* Complete entries \*

This entry is from: DDBJRELEASE:AE000154DDBJRELEASE**Save****Link****Launch****NCIustaIW** **Printer Friendly**

LOCUS AE000154 11519 bp DNA linear BCT 01-DEC-2000

DEFINITION Escherichia coli K12 MG1655 section 44 of 400 of the complete genome.

ACCESSION AE000154 U00096

VERSION AE000154.1

KEYWORDS .

SOURCE Escherichia coli K12

ORGANISM Escherichia coli K12

Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Escherichia.

REFERENCE 1 (bases 1 to 11519)

AUTHORS Blattner, F.R., Plunkett, G. III, Bloch, C.A., Perna, N.T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J.D., Rode, C.K., Mayhew, G.F., Gregor, J., Davis, N.W., Kirkpatrick, H.A., Goeden, M.A., Rose, D.J., Mau, B. and Shao, Y.

TITLE The complete genome sequence of Escherichia coli K-12

JOURNAL Science 277 (5331), 1453-1474 (1997)

MEDLINE 97426617

PUBMED 9278503

REFERENCE 2 (bases 1 to 11519)

AUTHORS Blattner, F.R.

TITLE Direct Submission

JOURNAL Submitted (16-JAN-1997) Guy Plunkett III, Laboratory of Genetics, University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA. Email: [ecoli@genetics.wisc.edu](mailto:ecoli@genetics.wisc.edu) Phone: 608-262-2534 Fax: 608-263-7459

REFERENCE 3 (bases 1 to 11519)

AUTHORS Blattner, F.R.

TITLE Direct Submission

JOURNAL Submitted (02-SEP-1997) Guy Plunkett III, Laboratory of Genetics, University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA. Email: [ecoli@genetics.wisc.edu](mailto:ecoli@genetics.wisc.edu) Phone: 608-262-2534 Fax: 608-263-7459

REFERENCE 4 (bases 1 to 11519)

AUTHORS Plunkett, G. III.

TITLE Direct Submission

JOURNAL Submitted (13-OCT-1998) Laboratory of Genetics, University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA

COMMENT This sequence was determined by the E. coli Genome Project at the University of Wisconsin-Madison (Frederick R. Blattner, director). Supported by NIH grants HG00301 and HG01428 (from the Human Genome Project and NCHGR). The entire sequence was independently determined from E. coli K12 strain MG1655. Predicted open reading frames were determined using GeneMark software, kindly supplied by Mark Borodovsky, Georgia Institute of Technology, Atlanta, GA, 30332 [e-mail: [mark@amherst.gatech.edu](mailto:mark@amherst.gatech.edu)]. Open reading frames that have been correlated with genetic loci are being annotated with CG Site Nos., unique ID nos. for the genes in the E. coli Genetic Stock Center (CGSC) database at Yale University, kindly supplied by Mary Berlyn. A public version of the database is accessible (<http://cgsc.biology.yale.edu>). Annotation of the genome is an ongoing task whose goal is to make the genome sequence more useful by correlating it with other data. Comments to the authors are appreciated. Updated information will be available at the E. coli Genome Project's World Wide Web site (<http://www.genetics.wisc.edu>). \*\*\* The E. coli K12 sequence and its annotations are periodically updated; this is version M54. No sequence changes. Annotation updates: updated gene identifications and products; all new functional assignments courtesy of Monica Riley; added promoters, protein binding sites, and repeated sequences described in reference 1. The unique numeric identifiers

beginning with a lowercase 'b' assigned to each gene (protein- or RNA-encoding) are now designated as gene synonyms instead of labels. This should allow them to be searched for in Entrez as gene names.

FEATURES	Location/Qualifiers
<u>source</u>	1..11519 /organism="Escherichia coli K12" /strain="K12" /sub_strain="MG1655" /db_xref="taxon:83333"
<u>promoter</u>	<1..26 /note="factor Sigma70; predicted +1 start at 499276"
<u>promoter</u>	complement(2..30) /note="factor Sigma70; predicted +1 start at 499238"
<u>gene</u>	106..1410 /gene="gsk" /note="synonym: b0477"
<u>CDS</u>	106..1410 /gene="gsk" /EC_number="2.7.1.73" /function="enzyme; Salvage of nucleosides and nucleotides" /note="o434; 100 pct identical to GB: ECOGSK_1 ACCESSION:D00798; 98 pct identical (1 gap) to INGK_ECOLI SW: P22937" /codon_start=1 /transl_table=11 /product="inosine-guanosine kinase" /protein_id="AAC73579.1" /db_xref="GI:1786684" /translation="MKFPGKRKSKHYFPVNARDPLLQQFQPENETSAAWVVGIDQTLV DIEAKVDDEFIERYGLSAGHSLVIEDDVAEALYQELKQKNLITHQFAGGTIGNTMHNY SVLADDRSVLLGVMCSNIEIGSYAYRYLCNTSSRTDLNLYLQGVDPGRGFTLIGESG ERTFAISPGHMNQLRAESIPEDVIAGASALVLTSLYLRCKPGPEPMPEATMKAIEYAKK YNVPVVLTLGTFVIAENPQWWQQFLKDHVSILAMNEDEAEALTGESDPLLASDKALD WVDLVLCAGPIGLYMAGFTEDEAKRKTQHPLPGAIAEFNQYEFSSRAMRHKDCQNPL RVYSHIAPYMGPEKIMNTNGAGDGALAALLHDITANSYHRSNVPNSSKHKFTWLTYS SLAQVCKYANRVSYQVLNQHSPLRGLPEREDSLEESYWDR"
<u>repeat region</u>	1425..1508 /note="REP (repetitive extragenic palindromic) element; contains 2 REP sequences"
<u>gene</u>	complement(1543..3219) /gene="ybaL" /note="synonym: b0478"
<u>CDS</u>	complement(1543..3219) /gene="ybaL" /function="putative transport; Not classified" /note="f558; 100 pct identical to fragment YBAL_ECOLI SW:P39830 (70 aa) but has 6 additional C-terminal aa and 485 additional N-terminal residues" /codon_start=1 /transl_table=11 /product="putative transport protein" /protein_id="AAC73580.1" /db_xref="GI:1786685" /translation="MHATPLITTIVGGLVLAFILGMLANKLRISPLVGYLLAGVLAG PFTPGFVADTKLAPELAELGVILLMFGVGLHFSKDLMAVKAIAIPGAIAQIAVATLL GMALSAVLGWSLMTGIVFGLCLSTASTVLLRALEERQLIDSQRGQIAIGWLIVEDLV MVLTLVLLPAVAGMMEQGDVGFATLAVDMGITIGKVIAFIAIMMLVGRRLVPWIMARS AATGSRELF TLSVLALALGVAFGAVELFVDFALGAFFAGMVLNESELSHRAAHTLP LRDAFAVLFFVSVGMLFDPLILIQQLAVLATLAILFGKSLAAFFLVRLFGHSQRTA LTIAASLAQIGEFALAGLGMALNLLPQAGQNLVLGAILSIMLNPVLFALLEKYLA KTETLEEQTLEEAIEEEKQIPVDICNHALLVGYGRVGSLLGEKLLASDIPLVVIETSR TRVDELRRGVRAVLGNAANEEIMQLAHLECAKWLILTIPNGYEAGEIVASARAKNPD IEIARAHYDDEVAYITERGANQVVMGEREIA RTMLELLETPPAGEVVTG"
<u>promoter</u>	complement(3328..3356) /note="factor Sigma70; predicted +1 start at 502564"
<u>gene</u>	complement(3457..4677) /gene="fsr" /note="synonym: b0479"
<u>CDS</u>	complement(3457..4677) /gene="fsr"

```

/function="putative transport; Not classified"
/note="f406; 100 pct identical to FSR_ECOLI SW: P52067"
/codon_start=1
/transl_table=11
/product="fosmidomycin resistance protein"
/protein_id="AAC73581.1"
/db_xref="GI:1786686"
/translation="MAMSEQPQPVAGAAASTTKARTSFGILGAISLSHLLNDMIQSLI
LA IYPLLQSEFSLTFMQIGMITLTFQLASSLLQPVVGYWTDKYPMPWSLPIGMCFTLS
GLVLLALAGSFGAVLLAAALVGTGSSVFHPESSRVARMASGGRHGLAQSIFQVGGNFG
SSLGPLLA AVI IAPYKGKGNVAVFVLAALLAIVVLAQISRWYSAQHRMNKGKPKATIIN
PLPRNKVVLAVSILLILIFSKYFYMASISSYYTFYLMQKFGLSIQNAQLHLFAFLFAV
AAGTVIGGPVGDKIGRKYVIWGSILGVAPFTLILPYASLHWTGVLTVIIGFILASAFS
AILVYAQELLPGRIGMVSGLFFGFAGMGGLGAAVLGLIADHTSIELVYKICAFPLPL
GMLTIFLPDNRHKD"
promoter complement(4711..4738)
/note="factor Sigma70; predicted +1 start at 503947"
promoter complement(4742..4770)
/note="factor Sigma70; predicted +1 start at 503978"
promoter complement(4770..4797)
/note="factor Sigma70; predicted +1 start at 504006"
promoter 4833..4861
/note="factor Sigma70; predicted +1 start at 504111"
promoter complement(4839..4865)
/note="factor Sigma70; predicted +1 start at 504075"
gene 4895..6547
/gene="ushA"
/note="synonym: b0480"
CDS 4895..6547
/gene="ushA"
/EC_number="3.6.1.45"
/function="enzyme; Central intermediary metabolism:
Sugar-nucleotide biosynthesis, conversions"
/note="o550; 99 pct identical to USHA_ECOLI SW: P07024"
/codon_start=1
/transl_table=11
/product="UDP-sugar hydrolase (5'-nucleotidase)"
/protein_id="AAC73582.1"
/db_xref="GI:1786687"
/translation="MKLLQRGVALALLTFTLASE TALAYEQDKTYK I TVLHTNDHHG
HFWRNEYGEYGLAAQKTLVDGIRKEVAAEGGSVLLS GGDINTGVPESDLQDAEPDFR
GMNLVGYDAMAI GNHEFDNPLTVLRQKEKWAKFPLLSANIYQKSTGERLFKPWALFKR
QDLKIAVIGLTTDDTAKIGNPEYFTDIEFRKPADEAKLVIQELQQTEKPDIIIAATHM
GHYDNGEHGSNAPGDVEMARALPAGSLAMIVGGHSQDPVCMAAENKKQVDYVPGT PCK
PDQQNGIWI VQAHEWGKYVGRADFEFRNGEMKMVNYQLIPVNLKKKVTWEDGKSERVL
YTPEIAENQQMISLLSPFQNKGAQLEVKI GETNGRLEGDRDKVRVQTNMGRLLAA
QMDRTGADFVMSGGGIRDSIEAGDISYKNVLKVQPFGNVVVYADMTGKEVIDYLTAV
AQMKPD SGAYPQFANVSFVAKDGKLNLDKIKGEPVDPAKTYRMATLNFNATGGDGYPR
LDNKPGYVNTGFIDAEVLKAYIQKSSPLDVS VYEPKGEVSWQ"
gene complement(6584..7063)
/gene="ybaK"
/note="synonym: b0481"
CDS complement(6584..7063)
/gene="ybaK"
/function="orf; Unknown"
/note="f159; 90 pct identical amino acid sequence and
equal length to YBAK_SALTY SW: P37174"
/codon_start=1
/transl_table=11
/product="orf, hypothetical protein"
/protein_id="AAC73583.1"
/db_xref="GI:1786688"
/translation="MTPAVKLLLEKNKISFQIHTYEHDP AETNFGDEVVKKLGLNPDQV
YKTL LVAVNGDMKHLAVAVTPVAGQLDLKKVAKALGAKKVEMADPMVAQRSTGYLVGG
ISPLGQKKRLPTIIDAPAQEFATIIYVSGGKRGDLIELAAGDLAKILD AKFADIARRD"
promoter complement(7211..7239)
/note="factor Sigma70; predicted +1 start at 506447"
gene complement(7267..8061)
/gene="ybaP"
/note="synonym: b0482"
CDS complement(7267..8061)

```

/gene="ybaP"  
/function="putative enzyme; Not classified"  
/note="f264; This 264 aa ORF is 25 pct identical (10 gaps)  
to 174 residues of an approx. 976 aa protein SLA2\_YEAST  
SW: P33338"  
/codon\_start=1  
/transl\_table=11  
/product="putative ligase"  
/protein\_id="AAC73584.1"  
/db\_xref="GI:1786689"  
/translation="MDLLYRVKTLWAALRGNHYTWPAIDITLPGNRHFHLIGSIHMGS  
HMAPLPTRLLKKLNADALIVEADVSTDPFANLPACEALEERISEEQLQNLQHIS  
QEMGISPSLFSTQPLWQIAMVLQATQAQKGLRAEYGYDYQLLQAAKQHKPVIELEG  
AENQIAMLLQLPDKGLALLDDTLTHWHTNARLLQQMMSWWLNAPPQNNDITLPNTFSQ  
SLYDVLMHQRNLAWRDKLRAMPPGRYVVAVGALHLYGEGNLPQMLR"  
promoter 7980..8007  
/note="factor Sigma70; predicted +1 start at 507257"  
promoter 8056..8085  
/note="factor Sigma70; predicted +1 start at 507335"  
gene 8145..8540  
/gene="ybaQ"  
/note="synonym: b0483"  
CDS 8145..8540  
/gene="ybaQ"  
/function="orf; Unknown"  
/note="o131; This 131 aa ORF is 38 pct identical (0 gaps)  
to 59 residues of an approx. 80 aa protein YSY1\_SYNP7 SW:  
P37371"  
/codon\_start=1  
/transl\_table=11  
/product="orf, hypothetical protein"  
/protein\_id="AAC73585.1"  
/db\_xref="GI:1786690"  
/translation="MIQYVLASLFTRKQQLKTMKQATRKPTTPGDILLYEYLEPLDLK  
INELAELLHVHRNSVSALINNNRKLTTMAFRLAKVFDTTVDVFWLNLQAAVDLWEVEN  
NMRTQEELGRIETVAEYLARREERAKKVA"  
promoter complement(8153..8179)  
/note="factor Sigma70; predicted +1 start at 507389"  
repeat\_region 8563..8798  
/note="REP (repetitive extragenic palindromic) element;  
contains 5 REP sequences"  
gene complement(8856..11360)  
/gene="ybaR"  
/note="synonym: b0484"  
CDS complement(8856..11360)  
/gene="ybaR"  
/function="putative enzyme; Not classified"  
/note="f834; This 834 aa ORF is 44 pct identical (16 gaps)  
to 735 residues of an approx. 752 aa protein ATCS\_SYNP7  
SW: P37279"  
/codon\_start=1  
/transl\_table=11  
/product="putative ATPase"  
/protein\_id="AAC73586.1"  
/db\_xref="GI:1786691"  
/translation="MSQTIDLTLDGLSCGHCVKRVKESLEQRPDVEQADVSITEAHVT  
GTASAEQLIETIKQAGYDASVSHPKAKPLAESSIPSEALTAVSEALPAATADDDDSQQ  
LLLSGMSASCVRVQNALQSVPGVTQARVNLAERTALVMGSASPQDLVQAVEKAGYG  
AEAIEDDAKRERQETAVATMKRFRWQAI VALAVGIPVMVWGMIGDNMMVTADNRSL  
WLVI GLITLAVMVFAGGHFYRSAWKSLLNGAATMDTLVALGTGVAWLYSMSVNLWPQW  
FPMEARHLYYEASAMIGLINLGHMLEARARQRSSKALEKLLDLPPTARLV TDEGEK  
SVPLAEVQPGMLLRLLTTGDRVPVDGEITQGEAWLDEAMLTGEP IPQQKGEGDSVHAGT  
VVQDGSVLFRAVGSHTTSLRIIRMVRAQSSKPEIGQLADKISAVFVPVVVIALV  
SAAIWYFFGPAPQIVYTLVIATTVLIIACPCALGLATPMSII SGVGRAAEFGVLVRDA  
DALQRASTLDTVFDKTGTLTEGKPQVVAVKTFADVDEAQLRLAAALEQGSSHPLAR  
A ILDKAGDMQLPQVNGFRTLRLGLGVSGEAGHALLGNQALLNEQQVGTKAIEAEITA  
QASQGATPVLLAVDGKAVALLAVRDLRSDSVAALQRLHKAGYRLVMLTGDNPPTANA  
IAKEAGIDEVIAGVLPDGKAEAIKHLQSEGRQVAMVGDGINAPALAQADVGIAMGGG  
SDVAIETAAITLMRHSLMGVADALAI SRATLHNMKNLLGAFIYNSIGIPVAAGILWP  
FTGTLLNPVVAGAAMALSSITVVSANRLLRFKPKPE"  
promoter complement(11401..11431)

```
/note="factor Sigma70; predicted +1 start at 510637"
promoter complement(11438..11466)
promoter /note="factor Sigma70; predicted +1 start at 510674"
11466..11495
promoter /note="factor Sigma70; predicted +1 start at 510745"
11504..>11519
promoter /note="factor Sigma70; predicted +1 start at 510782"

BASE COUNT      2883 a   3112 c   2969 g   2555 t
ORIGIN
1   gaaaatgcga tcccgccctgc tgatatgtgaa actggctgcg tctcgcgcg tcccgtcaga
61  ttgtgttaac attcgccgct cagttaacca cccgtaaaaa caaccatgaa atttcccgtt
121 aaacgtaaat ccaaacatta cttccccgta aacgcacgcg atccgctgct tcagcaattc
181 cagccagaaa acgaaaccag cgctgcctgg gtagtgggta tcgatcaaac gctggtcgat
241 attgaagcga aagtggatga tgaatttatt gagcgttatg gattaagcgc cgggcattca
301 ctggtgattg aggatgatgt agccgaagcg ctttatcagg aactaaaaca gaaaaacctg
361 attacccatc agtttgcggg tggcaccatt ggtaacacca tgcacaacta ctcggtgctc
421 gcggacgacc gttcggtgct gctgggcgtc atgtgcagca atattgaaat tggcagttat
481 gcctatcgtt acctgtgtaa cacttccagc cgtaccgatc ttaactatct acaaggcgtg
541 gatggcccga ttggctcgtt ctttacgttg attggcgagt ccggggaacg tacctttgct
601 atcagtccag gccacatgaa ccagctgcgg gctgaaagca ttccggaaga tgtgattgcc
661 ggagcctcgg cactggttct cacctcatat ctggtgcgtt gcaagccggg tgaacccatg
721 ccggaagcaa ccatgaaagc cattgagtag gcgaagaaat ataacgtacc ggtggtgctg
781 acgctgggca ccaagtttgt cattgccgag aatccgcagt ggtggcagca attcctcaaa
841 gatcacgtct ctatccttgc gatgaacgaa gatgaagccg aagcgttgac cggagaaagc
901 gatccgttgt tggcatctga caaggcgctg gactgggtag atctggtgct gtgcaccgcc
961 gggccaatcg gcttgtatat ggccggcctt accgaagacg aagcgaacg taaaaccag
1021 catccgctgc tgcggggcgc tatagcggaa ttcaaccagt atgagtttag ccgcgccatg
1081 cgccacaagg attgccagaa tccgctgcgt gtatatcgc acattgcgcc gtacatgggc
1141 gggccggaag aaatcatgaa cactaatgga gcgggggatg gcgcattggc agcgttgctg
1201 catgacatta ccgccaacag ctaccatcgt agcaacgtac caaactccag caaacataaa
1261 ttacacttgt taacttattc atcgttagcg cagggtgtga aatatgctaa ccgtgtgagc
1321 tatcaggtag tgaaccagca ttacactcgt ttaacgcgcg gcttgccgga gcgtgaagac
1381 agcctggaag agtcttactg ggatcgttaa gttatcgtcg gttcgtaggc cagataaggc
1441 gttcacgccc catctggcat ttggctctcg atgcctgatg cgacgctggc gcgtcttctc
1501 atgcctacat atttttcata ttttacatcc ggcaaccacc gtttaccggt tcaccacctc
1561 acccgccggt ggcgtttcca gcagttccag catggtacgg gcgatttcac gctcgcccat
1621 cactacctga ttgcgaccac gttcggtgat atacgccact tcatcgtcat aatgggcgcg
1681 ggcaataatc tcaatatccg gatttttcgc gcgggcagat gccacaatct caccgccttc
1741 ataaccgttg ggaatcgtca ggatcagcca ttttgcacat tccagatgag ccagttgcat
1801 aatttcttgc ttgcgcgcat tgcgaatac tgcgcggacc ccgcgctctc gcagctcttc
1861 aacacgggtt cgtgacgtct caatcaccac cagcgggaata tcagaggcga gcaatttctc
1921 cccagcagg ctgcctacac gaccgttaac caccagtagc gcatggttgc aaatatccac
1981 tgggatctgc ttctcttctt cgattgcctc ttccagcgtc tgctcttcca gcgtttcggt
2041 cttcgccaga tatttctcca gtagtgcgaa cagtaccggg ttgagcataa tcgacaggat
2101 cgccccctgc agtaccaggt tttgtccggc ctgcggcagt aaattcaatg ccattcccag
2161 tcccgccagg ataaacgcga actcaccaat ctgcgccagg ctggcggcga tggttaatgc
2221 cgtacgttgg gattgaccaa acagtcgcac caggaaaaat gcggctaacg acttaccaaa
2281 cagaataatc gccagcgtcg ccagcactgc cagcggttgc tgaatcagaa ttaacggatc
2341 aaacaacatc ccgacggaga caaaaaacag caccgcaaac gcgtcgcgca atggcagcgt
2401 atcgtgggcg gcacggtgac tcagttcaga ctggttcagt accatcccgg caaagaacgc
2461 accgagtgcg aaggagacat caaacagctc taccgcacca aaggcaacc ctaacgccag
2521 cgccagcacc gacagggtaa acagctcgcg agaaccggtt gccgcgctgc gtgccataat
2581 ccacggcacc agacggcgac ctaccagcat cataatggcg ataaatgcga tcactttgcc
2641 gatggtgatc cccatatcga ctgcaagagt ggcaaaagccc acatcgccct gttccatcat
2701 tcctgccact gcgggcagca acaccagcgt cagaaccatt accaggtctt ccacaatcaa
2761 ccaaccgatg gcgatttgc cagctgact gtcaattaat tgcgttctt caagtgcgcg
2821 cagtaacacc acggtactgg cggtggaag acataaacgg aacacgatac cggtcattaa
2881 cgaccagccc agcacggcag agagcgccat acccagcagc gtcgccacgg ctatctgggc
2941 gatcgccacc ggaatggcga tggcctttac cgccatcaaa tccttcagcg aaaagtgcga
3001 accgacgcca aacatcaaca gaatgacgcc cagttcagcc agttccgggg caagcttggg
3061 atcggcaaca aagcccggag tgaatggctc tgcagcaca cccgctaaca gatattccac
3121 cagaggagaa atacgtagt tattggccag catgccgagg ataaaggcga gcacaaggcc
3181 gccacaatg gtggtgataa gcgggtggc gtgatgcat cctctcctt ttcttgggtg
3241 ttattgtcca ttttggccg ggaaaaccaa aattacaggt aatagtttat gacaatttca
3301 ttgatgatgt tcatgaataa ttattgaatt ttgcagaaaa atggaattag ctgcaaaaaa
3361 agcacggatc ggaaaacaga aggcggttta acagaggaaa ggcttatggc gactgtaaaa
3421 cgatgcagcc aaagtttgcc tttggctgca atgaaatcag tctttatgcc ggttatcagg
3481 caggaatatg gtcaacatcc ccaatagtgg caggaaagca cagattttat agactaactc
3541 gatgctgggt tgcctggcga taagccccag aactgccgct cccagacctc ccatgccaaa
3601 agcaaaaccg aaaaagagtc cagaaacat accgatcgt cctggaagca gctcctgagc
3661 gtagaccaga atggcagaga atgccgaagc gaggataaat ccaataatca ccgttaaaac
```



3721 ccccggtccag tgcaggctgg cgtagggttaa aatcagcgta aacggcgcaa cgccgaggat  
3781 agagcccca atcacatatt tccgccaat tttatccctt acaggcccg cgatcaccgt  
3841 acctgccga acggcaaaca ggaaggcaaa cagatgaagc tgagcattct ggatagataa  
3901 tccgaatttt tgcatacagat aaaaggigta atagctgctg atgctcgcca tatagaaata  
3961 tttcgagaaa atgaggatta acagaaigct gaccgccagt acaactttat tgcgcggcag  
4021 tggattgata atcgtcgctt tgggttttcc tttattcatt cgggtgctgtg ccgagtlacca  
4081 acggctgatt tgcgccaaca ccacgatcgc cagcagtgcc gcaagcaca accaggcaac  
4141 gttgcctttg ccataaggcg cgataatcac cgccgccagc aagggtccca gggaactgcc  
4201 aaagtggcg ccgacctgaa agatagattg cgccaggcca tgcgccccg cggaagccat  
4261 acggggccag cgagaagatt ccggaigaaa gaccgatgaa ccggtaccga ccagcgccgc  
4321 cgccagcaga actgcgccaa aactgccccg cagcgcaagc agcaccagac cgcttaaggt  
4381 aaagcacatg ccaattggca acgacctgg catcgatat tttatcggtc agtagccgac  
4441 cactgggtgc agtagcgaag aggcgagctg gaaggtaggg gttatcatgc caatctgcat  
4501 aaatgtcaga gaaaattctg actgaagcag cggataaatc gccagaatca gcgattggat  
4561 catgtcgttc agcagatgtg agaggctgat agcacctaaa ataccaaacg atgttcgggc  
4621 ctgggtcggt gacgcagccg cgcccgccac aggcctgggt tggctactca ttgccatagg  
4681 aaagtcactt tttcagggtt gcgatgtaaa gaatgatctt atttggatt attaccagac  
4741 taacatacct gtatgcgtcg tctgaaggaa gtctcaacgc cgaatacaga atttctaate  
4801 tggatgcaga tttatcttca ccggacgcag acttgtctat gatgtcgctg catactatit  
4861 ttcaacacgt tgaatcagg tcaggagag agtatgaaa ttatgtagc gggcggtggc  
4921 gttagcgctg ttaaccacat ttacactggc gattgaaact gctctggcgt atgagcagga  
4981 taaaacctac aaaattacag ttctgcatac caatgatcat catgggcatt ttggcgcaa  
5041 tgaatatggc gaatatggc tggcgccgca aaaaacgctg gttgatggta tccgcaaaga  
5101 ggttgcggct gaaggcggtg gcgtgctgct actttccggt ggcgacatta acactggcgt  
5161 gcccgagctt gacttacagg atgccgaacc tgattttcgc ggtatgaatc tgggtggcta  
5221 tgacgcgatg gcgatcggtg atcatgaatt tgataatccg ctaccggtat tacgccagca  
5281 ggaaaagtgg gccaaagtcc cgttgctttc cgcaaatatc taccagaaaa gtactggcga  
5341 gcgcctgttt aaaccgtggg cgtgttttaa gcgtcaggat ctgaaaattg ccgttatagg  
5401 gctgacaacc gatgacacag caaaaattgg taaccgggaa tacttcactg atatcgaatt  
5461 tcgtaagccc gccgatgaag cgaagctggt gattcaggag ctgcaacaga cagaaaagcc  
5521 agacattatt atcgcgccga ccataatggg gcattacgat aatggtagag acggctctaa  
5581 cgcaccgggc gatgtggaga tggcacgcgc gctgcctgcc ggalcgctgg cgatgatcgt  
5641 cgggtggtcac tgcgaagatc cggctctgat ggcgccagaa aacaaaaaac aggtcgatta  
5701 cgtgccgggt acgccatgca aaccagatca aaaaaacggc atctggattg tgcaggcgca  
5761 tgagtggggc aaatacgtgg gacgggctga ttttgagttt cgtaatggcg aaatgaaat  
5821 ggttaactac cagctgattc cgttgaacct gaagaagaaa gtgacctggg aagacgggaa  
5881 aagcgagcgc gtgctttaca ctctgaaat cgtgaaaac cagcaaatga tctcgctgtt  
5941 atcaccgttc cagaacaaag gcaaagcgca gctggaagtg aaaaataggc aaaccaatgg  
6001 tcgtctggaa ggcatcgtg acaaagtgcg tttgtacag accaatatgg ggcggttgat  
6061 tctggcagcc caaatggatc gcactgggtg cgactttgcg gtgatgagcg gaggcggaat  
6121 tcgtgattct atcgaagcag gcgatatcag ctataaaaaac gtgctgaaag tgcagccatt  
6181 cggcaatgtg gtggtgtatg ccgacatgac cggtaaaagag gtgattgatt acctgaccgc  
6241 cgtcgcgtag atgaagccag attcagggtg ctaccgcaa tttgccacg tttagctttg  
6301 ggcgaaagac ggcaactga acgaccttaa aatcaaaggc gaaccggtcg atccggcgaa  
6361 aacttaccgt atggcgacat taaacttcaa tgccaccggc ggtgatggat atccgcgctt  
6421 tgataacaaa ccgggctatg tgaataccgg ctttattgat gccgaagtgc tgaagcgta  
6481 tatccagaaa agctcgccgc tggatgtgag tgtttatgaa ccgaaagggt aggtgagctg  
6541 gcagtaatcc gaaagtggcg gatgtttgca tccggcaca tgcttaatcg cggcggcgca  
6601 tatcagcaaa tttggcatcg aggatctttg ccagatcgcc tgcgcccagt tgcatalcca  
6661 gtccgcgctt gccgcccga aacataaatag tggcaaatc tttgtcgggg gcgtcgataa  
6721 tcgttggcag acgtttttc tgcgccagtg ggctaatccc cccaaccagg tatcccgctg  
6781 aacgctgcgc gaccatcgga tggccatct caactttctt ggcaaccagc gcttttgcta  
6841 cttttttaag atccagttga cctgcgaccg gcgtaacggc cacggcaagg tgtttcataat  
6901 caccgttcac tgccaccagc agcgttttgt agacctgatc cggattcaaa cctaattttt  
6961 tgacgacttc atcgccaaaa ttgggtttcag ccggatcgct ctcgtaggta tggatttgaa  
7021 acgaaatctt gtttttttcg agtaatttaa ctgcgggtgt catggtaat ttctgtcctt  
7081 aaacaaaaaa acgaagcaag catacgctg atggcggtcc aaaaaaaga gtcatcttgc  
7141 ctaagagtat tggcaggatg gtgagattga gcgacaatcg agttacaccg tgccttaag  
7201 tgacggcata ataataaaaa aatgaaatc ctctttgac ggccaatagc gatattggcc  
7261 atttttttag cgcaacattt gcggcaaat cctttctcca tacagggtga gtgcaccgac  
7321 cgcgaccaca tatcgccccg gcggcatggc gcgtaatita tcccgccagg cgagatttcg  
7381 ctgatgcata agcacatcgt acagcgactg actgaacgta ttgggcagcg ttatatcatt  
7441 attttgcggc ggtgcatcca gccaccagct catcatttgt tgcagcaacc gtgcgttgg  
7501 atgccagtg gtgcagctat cgtccagcag cgccagtcct ttgtcaggga gctggagcaa  
7561 catggcaatc tgggtttcag ccccttcag ttcaatcacg ggtttatgtt gttgcttcgc  
7621 cgcttgcaat agctggtaat cgataccgta ttctgcccgc agccccagtt tttgtgctg  
7681 cgtcgctga agaaccatcg cgatttgcca cagcggttgg gtagaaaaga gtgagggaga  
7741 aatgccatc tccggctaa tgtgctgcag gttttggagt tgcctctgc taatgcgctc  
7801 ttccagcgcc tgcagcgag gcaaatagc aaaaggcgta tgcgtgttgg aaacatccgc  
7861 ctgcagatc agcgcatcg cgtttttgag ctttttgagc aaacgggtgg gcaggggagc  
7921 catactgtg ctacccatat gaatactgcc aatcagatga aatggcgat tgcggggag

7981 ggtgatatcg atggcaggcc aggtgtaatg attaccgcgc agcgcggccc aaagcgtttt  
8041 tacccggtac aacagatcca tacgacctcc ctttgtgaaa tatcatgcta gcgcgcggtg  
8101 agggatggcg caaatggagt tggagacttg tttaatgtgt ttgtatgatt cagtatgttc  
8161 ttgcatcgct attcacaagg aagcaacagt taaaaacat gaaacaggca acaagaaaac  
8221 cgacgacacc cggagatatt ctctctatg aatatctgga accgctcgat ttgaaaatca  
8281 atgagttagc agagttgctg catgttcatc gtaatagcgt cagtgcactg atcaataaca  
8341 atcgtaaaact cactactgag atggcatttc gtcggcgaa agtttttgat accacagtcg  
8401 atttttggct aaacctccag gcgcgcggtg atctttggga agttgaaaac aacatgcgca  
8461 cccaggaaga attgggacgg atgaaacag tggctgaata ttggcacgc cgtgaagagc  
8521 gtgcaaaaaa ggtcgcgtaa ggcaaaaaca tcgcttcatc tcgtaggccg gataaggcgt  
8581 tcacgccgca tccggcatcc gagcaccatt gcctgatgcg acgcttaccg gtccttatcag  
8641 gcctacaaaac ttgtgcctga accgtaggcc ggataaggcg ttacgccgc atccggcatc  
8701 cgagcaccat tgcctgatgc gacgcttacg cgtcttatca ggctacaaa cctgtgcctg  
8761 aaccgtaggc cggataaggc gtttacgcgc catccggcag tcatgcgtcg atgcaaagt  
8821 cgccacccta aagcagcgca tccgcaatga tgtacttatt ccttcggttt aaaccgcagc  
8881 aaccggttgg cgttactcac tacggtaatc gacgagagcg ccatggccgc tccggcaact  
8941 accgggttaa gcagtgttc agtgaacggc caaaaatac cggcggcgac cggaataccg  
9001 atactgttgt agataaacgc accgagcagg ttctgcttca tgtgtgcag cgttgcgcgg  
9061 gaaatagcga gcgcatccgc aacgcccac aggctatggc gcatcagggt aatcgccgcg  
9121 gtttcaatgg caacatcact gccgccaccc atcgcaatgc cgacatccgc ctgagccagc  
9181 gctggcgctg cgttaatgcc gtgcgccacc attgccacct gacgtccttc actttgcaga  
9241 tgtttgatcg cttcggttt accgtccggc agcaccggcg cgatcacctc atcaatccct  
9301 gcttctttgg cgatcgcat tggcggtggt gggttatccc cggtaacat caccagacga  
9361 tatcccgttt tatgcaggcg ttgcagcgcc gccacgctat cactacgcaa cggatcgctg  
9421 actgccagca gggctaccgc ttcccgtca accgccagca gcacaggcgt tgcctttgc  
9481 gatgcctgag cagtaatctc cgcttcgata gctttgggtac caacctgttg ctcatttaac  
9541 agcgctgat tgcacagcaa taacgcatga ccttcagctt caccgctcac gccagcccg  
9601 cgcaatgtgc ggaaaccgtt gacctgcgtt agctgcata cacctgctt atcgaggatc  
9661 gctcgtgcca gcggatggct ggaaccttgc tccagtgcg ccgccagacg caatgcctgc  
9721 gcttcatcaa catcagcaaa tgttttact gcgacaacct gcggcttccc ttcagtcagc  
9781 gtcccggttt tatcgaacac tacagtgtcg agtgtactgg cgcgttgag cgcgtcagcg  
9841 tccgcacca gcacgcaaaa ctacgcccgc cggccagcg cggaataat cgacatcggc  
9901 gtgcagcacc ccagcgcaaa cggacaggca ataatacagta ccgtggtggc aatcaccagg  
9961 gtatagacaa tctgcggtgc cggacaaaag aaataccaga ttgccgact gacaagcgca  
10021 ataaccacca ctaccggcac aaatacggct gagattttat ccgccagctg accgatttct  
10081 ggcttgctgc tctgggcctg gcgcacctg cgaatgattc gtgacagcgt agtatggctg  
10141 ccaaccgcac tggcacgaaa cagcacactg ccgtcctgta ccactgtccc ggcatggacg  
10201 ctatcgcttt cgcctttttg ctgcgggatt ggttcgcccg tcagcatcgc ttcatccagc  
10261 catgcttcgc cctgggtaat ctgcacatct accggcacgc gatcgccggt cgtcaggcgc  
10321 agcaacatac ctggctgcac ttctgccaaga ggcacgcttt ttccacctc gtcagtaacc  
10381 aggcgtgccc tccgggggt taaatcgagt aacttttcca gcgccttaga agaacgctgg  
10441 cgtgcgcgcg cttccagcat atggccgaga ttgatcagac cgataatcat cgcgttggct  
10501 tcgtaataaaa gatgtcgcgc ttccatcggg aaccactgcg gccacagggt gacgctcatc  
10561 gaatagagcc acgccacgcc agtaccagc gccaccagcg tatccatcgt cgcgcacccg  
10621 ttcagcaggc ttttccatgc actgcggtaa aaatggccgc cggcgaaaac catcactgcc  
10681 agggttatca ggccgataac caaccacagg ctgcggttgt cagcggtgac catcatgta  
10741 tcgccgatca tccccagac catcaccggg ataccaccg ccagtgcgac aattgcctgc  
10801 cagcggagc gcttcatcgt agcgacggcg gtttcttgcg ggcgtcgcg gcgtttagcg  
10861 tcatcttcaa tcgcttccgc gccgtagccc gctttttcca ccgctgcac taaatcttgt  
10921 ggggaggcac tgcctatcac cagcgagta cgctccgcca ggtttaccg tgcctgagtg  
10981 acgcccggta cgttttgag cgcattttgt acgcggttga cacagctggc gcagctcatg  
11041 ccgctcagca gcaactgctg gctgtcatcg tcatcgccgg tcgctgccgg aagcgctca  
11101 gaaaccgctg tcagtgttcc cgacgggatt gatgactccg ccagcgggtt agcctttggg  
11161 tggcttacag atgcgtcata acccgcttgt ttgatggttt cgatcagctg ttctgcactg  
11221 gcagtcccgg taacgtgcgc ttcagtata gacacatccg cctgctcaac atccggacgc  
11281 tgttcaagac tttctttcac gcgtttaacg cagtaccgc aggacaggcc gtccagggtc  
11341 aggtcgatag ttgttgacat aaaacactcc ttaagacag ttttgactgg ctgtgataaa  
11401 ggttaaacct tccagcaagg ggaaggtaa gaaattaata aaccaggcgg gtaaaagtcc  
11461 gtaaagatta aaaaatcggc tcgatttgca tcaggattag acatttatct ctttgtttt

//